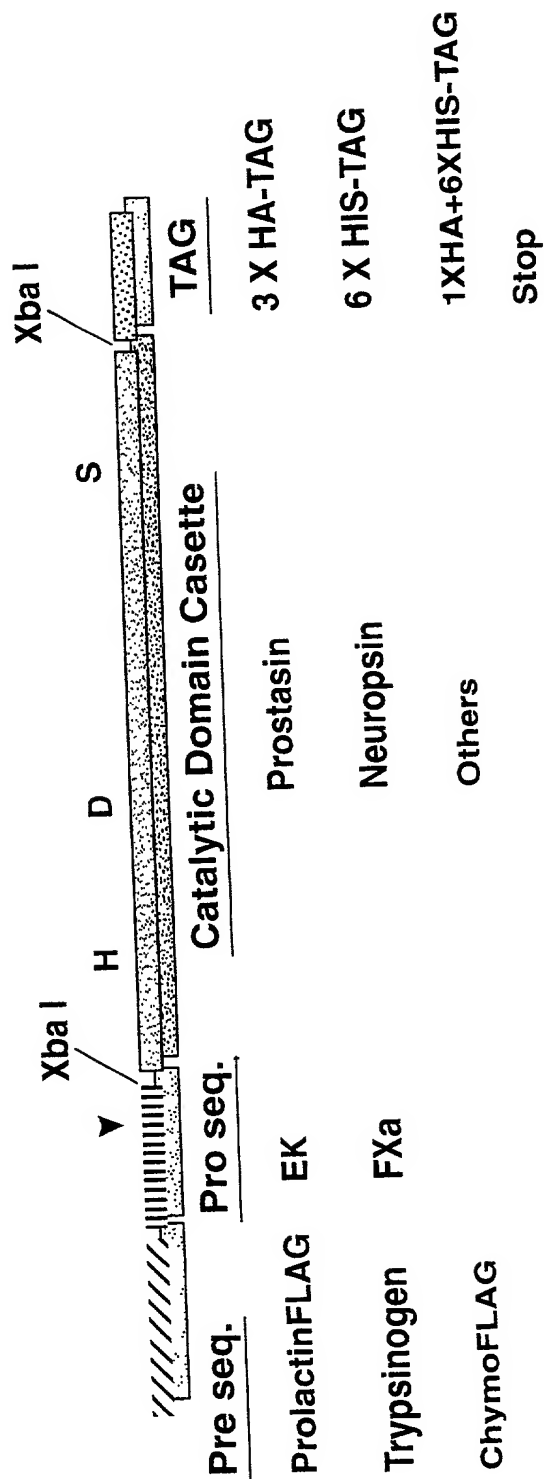


FIG. 1



Eco RI

GAATTCCACCACCATGGACAGCAAAGGTTTCGTTCGCAGAATCCCGCCTGCT
-----+-----+-----+-----+-----+
CTTAAGTGGTGGTACCTGTCTGTTCCAAGCAGCGTCTTTAGGGCGGACGA
M D S K G S S Q K S R L L
Prolactin Signal Sequence

CCTGCTGCTGGTGGTGTCAAATCTACTCTTGTCGCCAGGGTGTGGTCTCCG
-----+-----+-----+-----+-----+
GGACGACGACCACCACAGTTTAGATGAGAACACGGTCCCACACCAGAGGC
L L L V V S N L L L C Q G V V S
Prolactin Signal Sequence

Not I
ACTACAAGGACGACGACGACGTGGACGCGGCCGCTCTTGCTGCCCCCTTT
-----+-----+-----+-----+-----+
TGATGTTCTGCTGCTGCTGCACCTGCGCCGGCGAGAACGACGGGGGAAA
D Y K D D D D V D A A A L A A P F
FLAG EK2 Pro

Xba I Not I
GATGATGATGACAAGATCGTTGGGGGCTATGCTCTAGATAGCGGCCGCTT
-----+-----+-----+-----+-----+
CTACTACTACTGTTCTAGCAACCCCGATACGAGATCTATCGCCGGCGAA
D D D D K I V G G Y A L *
EK2 Pro

CCCTTTAGTGAGGGTTAATGCTTCGAGCAGACATGATAAGATACATTGAT
-----+-----+-----+-----+-----+
GGGAAATCACTCCCAATTACGAAGCTCGTCTGTACTATTCTATGTA ACTA
SV40 Late pA

GAGTTTGGACAAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTG
-----+-----+-----+-----+-----+
CTCAAACCTGTTTGGTGTGATCTTACGTCAC TTTTTTACGAAATAAAC
SV40 Late pA

TGAAATTTGTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATA
-----+-----+-----+-----+-----+
ACTTTAAACACTACGATAACGAAATAAACATTGGTAATATTCGACGTTAT
SV40 Late pA

HincII
AACAAAGTTGAC
-----+-----
TTGTTCAACTG

FIG. 2(B)

SEQ.ID.NO.:2

Eco RI Not I
 1 GAATTCACCATGAATCCACTCCTGATCCTTACCTTTGTGGCGGCCGCTCT 50
 -----+-----+-----+-----+-----+
 CTTAAGTGGTACTTAGGTGAGGACTAGGAATGGAAACACCGCCGGCGAGA
 M N P L L I L T F V A A A L
 Trypsinogen Pre

Xba I
 51 TGCTGCCCCCTTTGATGATGATGACAAGATCGTTGGGGGCTATTGTCTAG 100
 -----+-----+-----+-----+-----+
 ACGACGGGGGAACTACTACTACTGTTCTAGCAACCCCGATAACAGATC
 A A P F D D D D K I V G G Y C L
 EK3 Pro

Not I
 101 ATACCCCTACGATGTGCCCGATTACGCCTAGCGGCCGCTTCCCTTAGTG 150
 -----+-----+-----+-----+-----+
 TATGGGGATGCTACACGGGCTAATGCGGATCGCCGGCGAAGGGAAATCAC
 Y P Y D V P D Y A *
 1 X HA-TAG

151 AGGGTTAATGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGAC 200
 -----+-----+-----+-----+-----+
 TCCCAATTACGAAGCTCGTCTGTACTATTCTATGTAATACTCAAACCTG
 SV40 Late pA

201 AAACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGT 250
 -----+-----+-----+-----+-----+
 TTTGGTGTGATCTTACGTCAC TTTTACGAAATAAACACTTTAAACA
 SV40 Late pA

HincII
 251 GATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAACAAGTTGA 300
 -----+-----+-----+-----+-----+
 CTACGATAACGAAATAAACATTGGTAATATTCGACGTTATTTGTTCAACT
 SV40 Late

301 C 301
 -
 G

FIG. 2(C)

SEQ.ID.NO.:3

Eco RI
 1 GAATTCACCAACCATGGACAGCAAAGGTTCTGTCGCAGAAATCCCGCCTGCT 50
 -----+-----+-----+-----+-----+
 CTTAAGTGGTGGTACCTGTCTGTTTCCAAGCAGCGTCTTTAGGGCGGACGA
 M D S K G S S Q K S R L L
 Prolactin Signal Sequence

51 CCTGCTGCTGGTGGTGTCAAATCTACTCTTGTGCCAGGGTGTGGTCTCCG 100
 -----+-----+-----+-----+-----+
 GGACGACGACCACCACAGTTTAGATGAGAACACGGTCCCAACACCAGAGGC
 L L L V V S N L L L C Q G V V S
 Prolactin Signal Sequence

Not I
 101 ACTACAAGGACGACGACGACGTGGACGCGGCCGCTCTTGCTGCCCCCTTT 150
 -----+-----+-----+-----+-----+
 TGATGTTCTGCTGCTGCTGCACCTGCGCCGGCGAGAACGACGGGGGAAA
 D Y K D D D D V D A A A L A A P F
 FLAG FXa Pro

Xba I
 151 ATCGAGGGGCGCATTGTGGAGGGCTCGGATCTAGATACCCCTACGATGTG 200
 -----+-----+-----+-----+-----+
 TAGCTCCCCGCGTAACACCTCCCGAGCCTAGATCTATGGGGATGCTACAC
 I E G R I V E G S D L Y P Y D V
 FXa Pro

201 CCCGATTACGCCGCTAGATACCCCTACGATGTGCCCGATTACGCCGCTAG 250
 -----+-----+-----+-----+-----+
 GGGCTAATGCGGCGATCTATGGGGATGCTACACGGGCTAATGCGGCGATC
 P D Y A A R Y P Y D V P D Y A A R
 3 X HA-TAG

251 ATACCACTACGATGTGCCCGATTACGCCGCTAGATACCCCTACGATGTG 300
 -----+-----+-----+-----+-----+
 TATGGTGATGCTACACGGGCTAATGCGGCGATCTATGGGGATGCTACACG
 Y H Y D V P D Y A A R Y P Y D V
 3 X HA-TAG

Not I
 301 CCGATTACGCCTAGCGGCCGCTTCCCTTTAGTGAGGGTTAATGCTTCGAG 350
 -----+-----+-----+-----+-----+
 GGCTAATGCGGATCGCCGGCGAAGGGAAATCACTCCCAATTACGAAGCTC
 P D Y A *

FIG. 2(D)

351 CAGACATGATAAGATACATTGATGAGTTTGGACAAACCACA**ACT**AGAAATG 400
-----+-----+-----+-----+-----+
GTCTGTACTATTCTATGTA**ACT**ACTCAAACCTGTTTGGTGT**TGAT**CTTAC

SV40 Late pA

401 CAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTATT 450
-----+-----+-----+-----+-----+
GTCAC TTTT TACGAAATAAACACTTTAAACACTACGATAACGAAATAA

SV40 Late pA

HincII

451 TGTAACCATTATAAGCTGCAATAACAAGTTGAC 484
-----+-----+-----+-----
ACATTGGTAATATTCGACGTTATTTGTTCAACTG

FIG. 2(F)

HincII
 351 TAACCATTATAAGCTGCAATAAACAAAGTTGAC 382
 -----+-----+-----+--
 ATTGGTAATATTCGACGTTATTTGTTCAACTG

Eco RI

50

100

150

200

250

300

350

	II	
	AC	
351	--	352
	TG	

FIG. 2(H)

SEQ.ID.NO.:6

Eco RI
 1 GAATTCACCACCATGGCTTTCCTCTGGCTCCTCTCCTGCTGGGCCCTCCT 50
 -----+-----+-----+-----+-----+
 CTTAAGTGGTGGTACCGAAAGGAGACCGAGGAGAGGACGACCCGGGAGGA
 M A F L W L L S C W A L L
 Chymotrypsinogen Pre

51 GGGTACCACCTTCGGCTGCGGGGTCCCCGACTACAAGGACGACGACGACG 100
 -----+-----+-----+-----+-----+
 CCCATGGTGGGAAGCCGACGCCCCAGGGGCTGATGTTCTGCTGCTGCTGCTG
 G T T F G C G V P D Y K D D D D
 ---Chymotrypsinogen Pre-----FLAG-----

Not I
 101 CGGCCGCTCTTGCTGCCCCCTTTGATGATGATGACAAGATCGTTGGGGGC 150
 -----+-----+-----+-----+-----+
 GCCGGCGAGAACGACGGGGGAACTACTACTACTGTTCTAGCAACCCCG
 A A A L A A P F D D D D K I V G G
 -----EK2 Pro-----

Xba I
 151 TATGCTCTAGATACCCCTACGATGTGCCCGATTACGCCGCTAGACATCAC 200
 -----+-----+-----+-----+-----+
 ATACGAGATCTATGGGGATGCTACACGGGCTAATGCGGCGATCTGTAGTG
 Y A L Y P Y D V P D Y A A R H H
 HA 6 X HIS-TAG

Not I
 201 CATCACCATCACTAGCGGCCGCTTCCCTTTAGTGAGGGTTAATGCTTCGA 250
 -----+-----+-----+-----+-----+
 GTAGTGGTAGTGATCGCCGGCGAAGGGAAATCACTCCCAATTACGAAGCT
 H H H H *

251 GCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACAACCTAGAAT 300
 -----+-----+-----+-----+-----+
 CGTCTGTACTATTCTATGTAACACTCAAACCTGTTGGTGTTGATCTTA

SV40 Late pA

301 GCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTTAT 350
 -----+-----+-----+-----+-----+
 CGTCACTTTTTTTACGAAATAAACACTTTAAACACTACGATAACGAAATA

SV40 Late pA

10/34

FIG. 2(I)

351 HincII
TTGTAACCATTATAAGCTGCAATAACAAGTTGAC 385
-----+-----+-----+-----
AACATTGGTAATATTCGACGTTATTTGTTCAACTG

TTGTAACCATTATAAGCTGCAATAACAAGTTGAC

11/34

FIG. 3(A)

SEQ.ID.NO.:7

Eco RI
 1 GAATTCACCAACCATGGACAGCAAAGGTTTCGTCGCAGAAATCCCGCCTGCT 50
 -----+-----+-----+-----+-----+
 CTTAAGTGGTGGTACCTGTCTGTTTCCAAGCAGCGTCTTTAGGGCGGACGA
 M D S K G S S Q K S R L L
 Prolactin Signal Sequence

51 CCTGCTGCTGGTGGTGTCAAATCTACTCTTGTGCCAGGGTGTGGTCTCCG 100
 -----+-----+-----+-----+-----+
 GGACGACGACCACCACAGTTTAGATGAGAACACGGTCCCAACACCAGAGGC
 L L L V V S N L L L C Q G V V S
 Prolactin Signal Sequence

Not I
 101 ACTACAAGGACGACGACGACGTGGACGCGGCCGCTCTTGCTGCCCCCTTT 150
 -----+-----+-----+-----+-----+
 TGATGTTCTGCTGCTGCTGCACCTGCGCCGGCGAGAACGACGGGGGAAA
 D Y K D D D D V D A A A L A A P F
 FLAG EK2 Pro

Xba I
 151 GATGATGATGACAAGATCGTTGGGGGCTATGCTCTAGAGGCCGGTCAGTG 200
 -----+-----+-----+-----+-----+
 CTACTACTACTGTTCTAGCAACCCCGATACGAGATCTCCGGCCAGTCAC
 D D D D K I V G G Y A L E A G Q W
 EK2 Pro

201 GCCCTGGCAGGTCAGCATCACCTATGAAGGCGTCCATGTGTGTGGTGGCT 250
 -----+-----+-----+-----+-----+
 CGGGACCGTCCAGTCGTAGTGGATACTCCGCGAGGTACACACACCACCGA
 P W Q V S I T Y E G V H V C G G
 Protasin.CDS

251 CTCTCGTGTCTGAGCAGTGGGTGCTGTCAGCTGCTCACTGCTTCCCCAGC 300
 -----+-----+-----+-----+-----+
 GAGAGCACAGACTCGTCACCCACGACAGTCGACGAGTGACGAAGGGGTGCG
 S L V S E Q W V L S A A H C F P S
 Protasin.CDS

301 GAGCACCACAAGGAAGCCTATGAGGTCAAGCTGGGGGCCCAACAGCTAGA 350
 -----+-----+-----+-----+-----+
 CTCGTGGTGTTCCTTCGGATACTCCAGTTCGACCCCGGGTGGTCGATCT
 E H H K E A Y E V K L G A H Q L D
 Protasin.CDS

12/34

FIG. 3(B)

351 CTCTACTCCGAGGACGCCAAGGTCAGCACCCCTGAAGGACATCATCCCCC 400
-----+-----+-----+-----+-----+
GAGGATGAGGCTCCTGCGGTTCCAGTCGTGGGACTTCCTGTAGTAGGGGG
S Y S E D A K V S T L K D I I P H
----- Prostin.CDS -----

401 ACCCCAGCTACCTCCAGGAGGGCTCCCAGGGCGACATTGCACTCCTCCAA 450
-----+-----+-----+-----+-----+
TGGGGTCGATGGAGGTCTCCCGAGGGTCCCGCTGTAACTGAGGAGGTT
P S Y L Q E G S Q G D I A L L Q
----- Prostin.CDS -----

451 CTCAGCAGACCCATCACCTTCTCCCGCTACATCCGGGCCATCTGCCTCCC 500
-----+-----+-----+-----+-----+
GAGTCGTCTGGGTAGTGGGAAGAGGGCGATGTAGGCCGGGTAGACGGAGGG
L S R P I T F S R Y I R P I C L P
----- Prostin.CDS -----

501 TGCAGCCAACGCCTCCTTCCCCAACGGCCTCCACTGCACTGTCACTGGCT 550
-----+-----+-----+-----+-----+
ACGTCGGTTGCGGAGGAAGGGGTTGCCGGAGGTGACGTGACAGTGACCGA
A A N A S F P N G L H C T V T G
----- Prostin.CDS -----

551 GGGGTCATGTGGCCCCCTCAGTGAGCCTCCTGACGCCCAAGCCACTGCAG 600
-----+-----+-----+-----+-----+
CCCCAGTACACCGGGGGAGTCACTCGGAGGACTGCGGGTTCGGTGACGTC
W G H V A P S V S L L T P K P L Q
----- Prostin.CDS -----

601 CAACTCGAGGTGCCTCTGATCAGTCGTGAGACGTGTAAGTGCCTGTACAA 650
-----+-----+-----+-----+-----+
GTTGAGCTCCACGGAGACTAGTCAGCACTCTGCACATTGACGGACATGTT
Q L E V P L I S R E T C N C L Y N
----- Prostin.CDS -----

651 CATCGACGCCAAGCCTGAGGAGCCGCACTTTGTCCAAGAGGACATGGTGT 700
-----+-----+-----+-----+-----+
GTAGCTGCGGTTTCGGAATCCTCGGCGTGAAACAGGTTCTCCTGTACCACA
I D A K P E E P H F V Q E D M V
----- Prostin.CDS -----

FIG. 3(C)

701 GTGCTGGCTATGTGGAGGGGGGCAAGGACGCCTGCCAGGGT GACTCTGGG 750
 -----+-----+-----+-----+-----+
 CACGACCGATACACCTCCCCCGTTCTGCGGACGGTCCCACTGAGACCC
 C A G Y V E G G K D A C Q G D S G
 ----- Prostin.CDS -----

751 GGCCCACTCTCCTGCCCTGTGGAGGGTCTCTGGTACCTGAC GGGCATTGT 800
 -----+-----+-----+-----+-----+
 CCGGGTGAGAGGACGGGACACCTCCAGAGACCATGGACTGCCCGTAACA
 G P L S C P V E G L W Y L T G I V
 ----- Prostin.CDS -----

801 GAGCTGGGGAGATGCCTGTGGGGCCCGCAACAGGCCTGGTGTGTACACTC 850
 -----+-----+-----+-----+-----+
 CTCGACCCCTCTACGGACACCCCGGGCGTTGTCCGGACCACACATGTGAG
 S W G D A C G A R N R P G V Y T
 ----- Prostin.CDS -----

851 TGGCCTCCAGCTATGCCTCCTGGATCCAAAGCAAGGTGACA GAACTCCAG 900
 -----+-----+-----+-----+-----+
 ACCGGAGGTTCGATACGGAGGACCTAGGTTTCGTTCCACTGTCTTGAGGTC
 L A S S Y A S W I Q S K V T E L Q
 ----- Prostin.CDS -----

901 CCTCGTGTGGTGCCCCAAACCCAGGAGTCCAGCCCGACAGCAACCTCTG 950
 -----+-----+-----+-----+-----+
 GGAGCACACCACGGGGTTTGGGTCCTCAGGGTTCGGGCTGTCTGTTGGAGAC
 P R V V P Q T Q E S Q P D S N L C
 ----- Prostin.CDS -----

951 TGGCAGCCACCTGGCCTTCAGCTCTAGACATCACCATCACCATCACTAGC 1000
 -----+-----+-----+-----+-----+
 ACCGTCGGTGGACCGGAAGTCGAGATCTGTAGTGGTAGTG GTAGTGATCG
 G S H L A F S | S R | H H H H H H *
 ----- Prostin.CDS ----- 6 X HIS-TAG -----

1001 Not I
 GGCCGCTTCCCTTTAGTGAGGGTTAATGCTTCGAGCAGACATGATAAGAT 1050
 -----+-----+-----+-----+-----+
 CCGGCGAAGGGAAATCACTCCCAATTACGAAGCTCGTCTGTACTATTCTA

FIG. 3(D)

1051 ACATTGATGAGTTTGGACAAACCACAACCTAGAATGCAGTGA~~AAAAAATGC~~ 1100
 -----+-----+-----+-----+
 TGTAACACTCAAACCTGTTTGGTGTGATCTTACGTCAC~~T TTTTACG~~

SV40 Late pA

1101 TTTATTTGTGAAATTTGTGATGCTATTGCTTTATTTGTAAC~~CATTATAAG~~ 1150
 -----+-----+-----+-----+
 AAATAAACACTTTAAACACTACGATAACGAAATAAACATTG~~GTAATATTC~~

SV40 Late pA

1151 CTGCAATAAACAAGTTGAC 1169
 -----+-----
 GACGTTATTTGTTCAACTG

15/34

FIG. 4(A)

SEQ.ID.NO.:8

Eco RI
 1 GAATTCACCATGGCTTTCTCTGGCTCCTCTCCTGCTGGGCCCCTCCT 50
 -----+-----+-----+-----+-----+
 CTTAAGTGGTGGTACCGAAAGGAGACCGAGGAGAGGACGACCCGGGAGGA
 M A F L W L L S C W A L L
 Chymotrypsinogen Pre

51 GGTACCACCTTCGGCTGCGGGGTCCCCGACTACAAGGACGACGACGACG 100
 -----+-----+-----+-----+-----+
 CCCATGGTGAAGCCGACGCCCCAGGGGGCTGATGTTCTGCTGCTGCTGC
 G T T F G C G V P D Y K D D D D
 Chymotrypsinogen Pre FLAG

Not I
 101 CGGCCGCTCTTGCTGCCCCCTTTGATGATGATGACAAGATCGTTGGGGGC 150
 -----+-----+-----+-----+-----+
 GCCGGCGAGAACGACGGGGGAACTACTACTGTTCTAGCAACCCCCG
 A A A L A A P F D D D D K I V G G
 EK2 Pro

Xba I
 151 TATGCTCTAGAGCCGGTCAGTGGCCCTGGCAGGTCAGCATCACCTATGA 200
 -----+-----+-----+-----+-----+
 ATACGAGATCTCCGGCCAGTCACCGGGACCGTCCAGTCGTAGTGGATACT
 Y A L E A G Q W P W Q V S I T Y E
 Prostasin.CDS

201 AGGCGTCCATGTGTGTGGTGGCTCTCTCGTGTCTGAGCAGTGGGTGCTGT 250
 -----+-----+-----+-----+-----+
 TCCGCAGGTACACACACCACCGAGAGAGCACAGACTCGTCACCCACGACA
 G V H V C G G S L V S E Q W V L
 Prostasin.CDS

251 CAGCTGCTCACTGCTTCCCCAGCGAGCACCACAAGGAAGCCTATGAGGTC 300
 -----+-----+-----+-----+-----+
 GTCGACGAGTGACGAAGGGGTGCTCGTGGTGTTCCTTCGGATACTCCAG
 S A A H C F P S E H H K E A Y E V
 Prostasin.CDS

301 AAGCTGGGGGCCACCAGCTAGACTCCTACTCCGAGGACGCCAAGGTCAG 350
 -----+-----+-----+-----+-----+
 TTCGACCCCCGGGTGGTCTGATCTGAGGATGAGGCTCCTGCGGTTCCAGTC
 K L G A H Q L D S Y S E D A K V S
 Prostasin.CDS

351	CACCCGTGAAGGACATCATCCCCCACCAGCTACCTCCAGGAGGGGCTCCC -----+-----+-----+-----+-----+ GTGGGACTTCCTGTAGTAGGGGGTGGGGTCGATGGAGGTCC TCCCGAGGG T L K D I I P H P S Y L Q E G S ----- Prostasin.CDS -----	400
401	AGGGCGACATTGCACTCCTCCAACCTCAGCAGACCCATCACCTTCTCCCGC -----+-----+-----+-----+-----+ TCCCGCTGTAACGTGAGGAGGTTGAGTCGTCTGGGTAGTGGAAGAGGGCG Q G D I A L L Q L S R P I T F S R ----- Prostasin.CDS -----	450
451	TACATCCGGCCCATCTGCCTCCCTGCAGCCAACGCCTCCTTCCCCAACGG -----+-----+-----+-----+-----+ ATGTAGGCCGGGTAGACGGAGGGACGTCGGTTGCGGAGGAAGGGGTTGCC Y I R P I C L P A A N A S F P N G ----- Prostasin.CDS -----	500
501	CCTCCACTGCACTGTCACTGGCTGGGGTCATGTGGCCCCCTCAGTGAGCC -----+-----+-----+-----+-----+ GGAGGTGACGTGACAGTGACCGACCCCAGTACACCGGGGGAGTCACTCGG L H C T V T G W G H V A P S V S ----- Prostasin.CDS -----	550
551	TCCTGACGCCCAAGCCACTGCAGCAACTCGAGGTGCCTCTGATCAGTCGT -----+-----+-----+-----+-----+ AGGACTGCGGGTTTCGGTGACGTCGTTGAGCTCCACGGAGACTAGTCAGCA L L T P K P L Q Q L E V P L I S R ----- Prostasin.CDS -----	600
601	GAGACGTGTAACCTGCCTGTACAACATCGACGCCAAGCCTGAGGAGCCGCA -----+-----+-----+-----+-----+ CTCTGCACATTGACGGACATGTTGTAGCTGCGGTTGCGACTCCTCGGCGT E T C N C L Y N I D A K P E E P H ----- Prostasin.CDS -----	650
651	CTTTGTCCAAGAGGACATGGTGTGTGCTGGCTATGTGGAGGGGGGGCAAGG -----+-----+-----+-----+-----+ GAAACAGGTTCTCCTGTACCACACACGACCGATAACCTCCCCCGTTCC F V Q E D M V C A G Y V E G G K ----- Prostasin.CDS -----	700

17/34

FIG. 4(c)

701 ACGCCTGCCAGGGTGA CTCTGGGGGCCCACTCTCCTGCCCTGTGGAGGGT 750
 -----+-----+-----+-----+-----+
 TCGGGACGGTCCCAC TGAACCCCCGGGTGAGAGGACGGGACACCTCCA
 D A C Q G D S G G P L S C P V E G
 -----Prostasin.CDS-----

751 CTCTGGTACCTGACGGGCATTGTGAGCTGGGGAGATGCCTGTGGGGCCCG 800
 -----+-----+-----+-----+-----+
 GAGACCATGGACTGCCCGTAACACTCGACCCCTCTACGGACACCCCGGGC
 L W Y L T G I V S W G D A C G A R
 -----Prostasin.CDS-----

801 CAACAGGCCTGGTGTGTACACTCTGGCCTCCAGCTATGCCTCCTGGATCC 850
 -----+-----+-----+-----+-----+
 GTTGTCCGGACCAACACATGTGAGACCGGAGGTGATACGGAGGACCTAGG
 N R P G V Y T L A S S Y A S W I
 -----Prostasin.CDS-----

851 AAAGCAAGGTGACAGAACTCCAGCCTCGTGTGGTGCCCCAAACCCAGGAG 900
 -----+-----+-----+-----+-----+
 TTTCTGTTCCACTGTCTTGAGGTGCGAGCACACCACGGGGTTTGGGTCTC
 Q S K V T E L Q P R V V P Q T Q E
 -----Prostasin.CDS-----

901 TCCCAGCCCGACAGCAACCTCTGTGGCAGCCACCTGGCCTTCAGCTCTAG 950
 -----+-----+-----+-----+-----+
 AGGGTCGGGCTGTCTGTTGGAGACACCGTGGTGGACCGGAAGTTCGAGATC
 S Q P D S N L C G S H L A F S S R
 -----Prostasin.CDS-----

951 ACATCACCATCACCATCACTAGCGGCCGCTTCCCTTTAGTGAGGGTTAAT 1000
 -----+-----+-----+-----+-----+
 TG TAGTGGTAGTGGTAGTGATCGCCGGCGAAGGGAAATCACTCCCAATTA
 | H H H H H H * |
 ----- 6 X HIS-TAG -----

1001 GCTTCGAGCAGACATGATAAGATACATTGATGAGTTGGACAAACCACAA 1050
 -----+-----+-----+-----+-----+
 CGAAGCTCGTCTGTACTATTCTATGTAAC TACTCAAACCTGTTTGGTGTT

SV40 Late pA

FIG. 4(D)

1051 CTAGAATGCAGTGAAAAAATGCTTATTTGTGAAATTGTGATGCTATT 1100
-----+-----+-----+-----+-----+
GATCTTACGTCACCTTTTTTTACGAAATAAACACTTTAAACA CTACGATAA
SV40 Late pA

1101 GCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTGAC 1142
-----+-----+-----+-----+-----
CGAAATAAACATTGGTAATATTCGACGTTATTTGTTCAACTG

19/34

FIG. 5(A)

SEQ.ID.NO.:9

Eco RI
1 GAATTCACCACCATGGACAGCAAAGGTTCGTCGCAGAAATCCCGCTGCT 50
-----+-----+-----+-----+-----+
CTTAAGTGGTGGTACCTGTCGTTTCCAAGCAGCGTCTTTAGGGCGGACGA
M D S K G S S Q K S R L L
Prolactin Signal Sequence

51 CCTGCTGCTGGTGGTGTCAAATCTACTCTTGTGCCAGGGTGTGGTCTCCG 100
-----+-----+-----+-----+-----+
GGACGACGACCACCACAGTTTAGATGAGAACACGGTCCCACACCAGAGGC
L L L V V S N L L L C Q G V V S
Prolactin Signal Sequence

Not I
101 ACTACAAGGACGACGACGACGTGGACGCGCCGCTCTTGCTGCCCCCTTT 150
-----+-----+-----+-----+-----+
TGATGTTCTGCTGCTGCTGCACCTGCGCCGGCGAGAACGACGGGGGAAA
D Y K D D D D V D A A A L A A P F
FLAG EK1 Pro

Xba I
151 GATGATGATGACAAGATCGTTGGGGGCTACAACCTGTCTAGAACCCCATTC 200
-----+-----+-----+-----+-----+
CTACTACTACTGTTCTAGCAACCCCCGATGTTGACAGATCTTGGGGTAAG
D D D D K I V G G Y N C L E P H S
EK1 Pro

201 GCAGCCTTGGCAGGCGGCCTTGTTCAGGGCCAGCAACTACTCTGTGGCG 250
-----+-----+-----+-----+-----+
CGTCGGAACCGTCCGCCGGAACAAGGTCCCGGTGCGTTGATGAGACACCGC
Q P W Q A A L F Q G Q Q L L C G
Neuropsin.CDS

251 GTGTCCTTGTAGGTGGCAACTGGGTCCTTACAGCTGCCCCACTGTAAAAAA 300
-----+-----+-----+-----+-----+
CACAGGAACATCCACCGTTGACCCAGGAATGTCGACGGGTGACATTTTTT
G V L V G G N W V L T A A H C K K
Neuropsin.CDS

301 CCGAAATACACAGTACGCCTGGGAGACCACAGCCTACAGAAATAAAGATGG 350
-----+-----+-----+-----+-----+
GGCTTTATGTGTCATGCGGACCCTCTGGTGTGCGATGTCTTATTCTACC
P K Y T V R L G D H S L Q N K D G
Neuropsin.CDS

FIG. 5(B)

351 CCCAGAGCAAGAAATACCTGTGGTTCAGTCCATCCCACACCCTGCTACA 400
 -----+-----+-----+-----+-----+
 GGGTCTCGTTCTTTATGGACACCAAGTCAGGTAGGGTGTGGGGACGATGT
 P E Q E I P V V Q S I P H P C Y
 ----- Neuropsin.CDS -----

401 ACAGCAGCGATGTGGAGGACCACAACCATGATCTGATGCTTCTTCAACTG 450
 -----+-----+-----+-----+-----+
 TGTCGTCGCTACACCTCCTGGTGTGGTACTAGACTACGAAGAAGTTGAC
 N S S D V E D H N H D L M L L Q L
 ----- Neuropsin.CDS -----

451 CGTGACCAGGCATCCCTGGGGTCCAAAGTGAAGCCCATCAGCCTGGCAGA 500
 -----+-----+-----+-----+-----+
 GCACTGGTCCGTAGGGACCCAGGTTTCACTTCGGGTAGTCGGACCGTCT
 R D Q A S L G S K V K P I S L A D
 ----- Neuropsin.CDS -----

501 TCATTGCACCCAGCCTGGCCAGAAGTGCACCGTCTCAGGCTGGGGCACTG 550
 -----+-----+-----+-----+-----+
 AGTAACGTGGGTGCGGACCGGTCTTCACGTGGCAGAGTCCGACCCCGTGAC
 H C T Q P G Q K C T V S G W G T
 ----- Neuropsin.CDS -----

551 TCACCAGTCCCCGAGAGAATTTTCTGACACTCTCAACTGTGCAGAAGTA 600
 -----+-----+-----+-----+-----+
 AGTGGTCAGGGGCTCTCTTAAAGGACTGTGAGAGTTGACAAGTCTTCAT
 V T S P R E N F P D T L N C A E V
 ----- Neuropsin.CDS -----

601 AAAATCTTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGGCAGATCAC 650
 -----+-----+-----+-----+-----+
 TTTTAGAAAGGGGTCTTCTTCACACTCCTACGAATGGGCCC CGTCTAGTG
 K I F P Q K K C E D A Y P G Q I T
 ----- Neuropsin.CDS -----

651 AGATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGGCTGACACGTGCCAGG 700
 -----+-----+-----+-----+-----+
 TCTACCGTACCAGACACGTCCGTCGTCGTTTCCCCGACTGTGCACGGTCC
 D G M V C A G S S K G A D T C Q
 ----- Neuropsin.CDS -----

FIG. 5(C)

701 GCGATTCTGGAGGCCCCCTGGTGTGTGATGGTGCCTCCAGGGGCATCACA 750
 -----+-----+-----+-----+-----+
 CGCTAAGACCTCCGGGGGACCACACACTACCACGTGAGGTCCCGTAGTGT
 G D S G G P L V C D G A L Q G I T
 ----- Neuropsin.CDS -----

751 TCCTGGGGCTCAGACCCCTGTGGGAGGTCCGACAAACCTGGCGTCTATAC 800
 -----+-----+-----+-----+-----+
 AGGACCCCGAGTCTGGGGACACCCTCCAGGCTGTTTGGACCGCAGATATG
 S W G S D P C G R S D K P G V Y T
 ----- Neuropsin.CDS -----

801 CAACATCTGCCGCTACCTGGACTGGATCAAGAAGATCATAGGCGACGAAGG 850
 -----+-----+-----+-----+-----+
 GTTGTAGACGGCGATGGACCTGACCTAGTTCTTCTAGTATCCGTCGTTCC
 N I C R Y L D W I K K I I G S K
 ----- Neuropsin.CDS -----

851 Xba I Not I
 GCTCTAGACATCACCATCACCATCACTAGCGGCCGCTCCCTTTAGTGAG 900
 -----+-----+-----+-----+-----+
 CGAGATCTGTAGTGGTAGTGGTAGTGATCGCCGGCGAAGGGAAATCACTC
 G | S R | H H H H H H *
 ----- 6 X HIS-TAG -----

901 GGTAAATGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGGACAA 950
 -----+-----+-----+-----+-----+
 CCAATTACGAAGCTCGTCTGTACTATTCTATGTAATACTCAAACTGTT

 SV40 Late pA

951 ACCACAAC TAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTGTGA 1000
 -----+-----+-----+-----+-----+
 TGGTGTTGATCTTACGTCACTTTTTTTACGAAATAAACACTTTAAACT

 SV40 Late pA

1001 TGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAAGTTGAC 1049
 -----+-----+-----+-----+-----+
 ACGATAACGAAATAAACATTGGTAATATTCGACGTTATTTGTTCAACTG

 SV40 Late pA

Eco RI

CCTGCTGCTGGTGGTGTCAAATCTACTCTTGCGCCAGGGT**G**TGGTTCCG
-----+-----+-----+-----+-----+
GGACGACGACCACCACAGTTTTAGATGAGAACACGGTCCCACA**ACC**AGAGGC
L L L V V S N L L L C Q G V V S |
Prolactin Signal Sequence

100

Xba I

151 GATGATGATGACAAGATCGTTGGGGGCTACAACTGTCTAGA AAAGCACTC
-----+-----+-----+-----+-----+-----+-----+
CTACTACTACTGTTCTAGCAACCCCGATGTTGACAGATCT TTTTCGTGAG
D D D D K I V G G Y N C L E K H S
-----EKL Pro-----

200

251 CGACGCTCATCGCCCCAGATGGCTCCTGACAGCAGCCCCACTGCCTCAAG 300
-----+-----+-----+-----+-----+
GCTGCGAGTAGCGGGGGTCTACCGAGGACTGTCGTCGGGTGACGGAGTTC
A T L I A P R W L L T A A H C L K
-----Protease O.CDS-----

301 CCCCCTACATAGTTCACCTGGGGCAGCACAACTCCAGAAAGGAGGAGGG 350
-----+-----+-----+-----+
GGGGCGATGTATCAAGTGGACCCCGTCGTGTTGGAGGTCTTCTCCTCTCCC
P R Y I V H L G Q H N L Q K E E G
-----Protease O.CDS-----

351	CTGTGAGCAGACCCGGACAGCCACTGAGTCCTTCCCCACCCCGGCTTCA -----+-----+-----+-----+-----+ GACACTCGTCTGGGCCTGTCGGTGACTCAGGAAGGGGGTGGGGCCGAAGT C E Q T R T A T E S F P H P G F -----Protease 0.CDS-----	400
401	ACAACAGCCTCCCCAACAAAGACCACCGCAATGACATCATGCTGGTGAAG -----+-----+-----+-----+-----+ TGTTGTGCGGAGGGGTTGTTTCTGGTGGCGTTACTGTAGTACGACCACTTC N N S L P N K D H R N D I M L V K -----Protease 0.CDS-----	450
451	ATGGCATCGCCAGTCTCCATCACCTGGGCTGTGCGACCCCTCACCCCTCTC -----+-----+-----+-----+-----+ TACCGTAGCGGTGAGAGGTAGTGGACCCGACACGCTGGGGAGTGGGAGAG M A S P V S I T W A V R P L T L S -----Protease 0.CDS-----	500
501	CTCACGCTGTGTCACTGCTGGCACCAGCTGCCTCATTTCCGGCTGGGGCA -----+-----+-----+-----+-----+ GAGTGCACACAGTGACGACCGTGGTGCACGGAGTAAAGGCCGACCCCGT S R C V T A G T S C L I S G W G -----Protease 0.CDS-----	550
551	GCACGTCCAGCCCCAGTTACGCCTGCCTCACACCTTGGGATGCGCCAAC -----+-----+-----+-----+-----+ CGTGCAGGTGCGGGGTCAATGCGGACGGAGTGTGGAACGTACGCGGTTG S T S S P Q L R L P H T L R C A N -----Protease 0.CDS-----	600
601	ATCACCATCATTGAGCACCAGAAGTGTGAGAACGCCTACCCCGGCAACAT -----+-----+-----+-----+-----+ TAGTGGTAGTAACCTCGTGGTCTTCACACTCTTGCGGATGGGGCCGTTGTA I T I I E H Q K C E N A Y P G N I -----Protease 0.CDS-----	650
651	CACAGACACCATGGTGTGTGCCAGCGTGCAGGAAGGGGGCAAGGACTCCT -----+-----+-----+-----+-----+ GTGTCTGTGGTACCACACACGGTGCACGTCCTTCCCCCGTTCTCTGAGGA T D T M V C A S V Q E G G K D S -----Protease 0.CDS-----	700

FIG. 6(C)

701 GCCAGGGTGACTCCGGGGGCCCTCTGGTCTGTAACCACTCTCTTCAAGGC 750
 -----+-----+-----+-----+-----+
 CGGTCCCCTGAGGCCCCCGGGAGACCAGACATTGGTCAGAGAAGTTCCG
 C Q G D S G G P L V C N Q S L Q G
 -----Protease O.CDS-----

751 ATTATCTCCTGGGGCCAGGATCCGTGTGCGATCACCCGAAAGCCTGGTGT 800
 -----+-----+-----+-----+-----+
 TAATAGAGGACCCCGGTCTAGGCACACGCTAGTGGGCTTTTCGGACCACA
 I I S W G Q D P C A I T R K P G V
 -----Protease O.CDS-----

801 CTACACGAAAGTCTGCAAATATGTGGACTGGATCCAGGAGACGATGAAGA 850
 -----+-----+-----+-----+-----+
 GATGTGCTTTTCAGACGTTTATACACCTGACCTAGGTCCTCTGCTACTTCT
 Y T K V C K Y V D W I Q E T M K
 -----Protease O.CDS-----

851 Xba I Not I
 ACAATTCTAGACATCACCATCACCATCACTAGCGGCCGCTTCCCTTTAGT 900
 -----+-----+-----+-----+-----+
 TGTTAAGATCTGTAGTGGTAGTGGTAGTGATCGCCGGCGAAGGGGAATCA
 N N | S R | H H H H H H *
 -----6 X HIS-TAG-----

901 GAGGGTTAATGCTTCGAGCAGACATGATAAGATACATTGATGAGTTTGA 950
 -----+-----+-----+-----+-----+
 CTCCCAATTACGAAGCTCGTCTGTACTATTCTATGTAACCTACTCAACCT

 SV40 Late pA

951 CAAACCACAACTAGAAATGCAGTGAAAAAATGCTTTATTTGTGAAATTTG 1000
 -----+-----+-----+-----+-----+
 GTTTGGTGTGATCTTACGTCACCTTTTTTTACGAAATAAACACTTTAAAC

 SV40 Late pA

1001 TGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAAGTTG 1050
 -----+-----+-----+-----+-----+
 ACTACGATAACGAAATAAACATTGGTAATATTGACGTTATTTGTTCAAC

 SV40 Late pA

1051 AC
 -- 1052
 TG

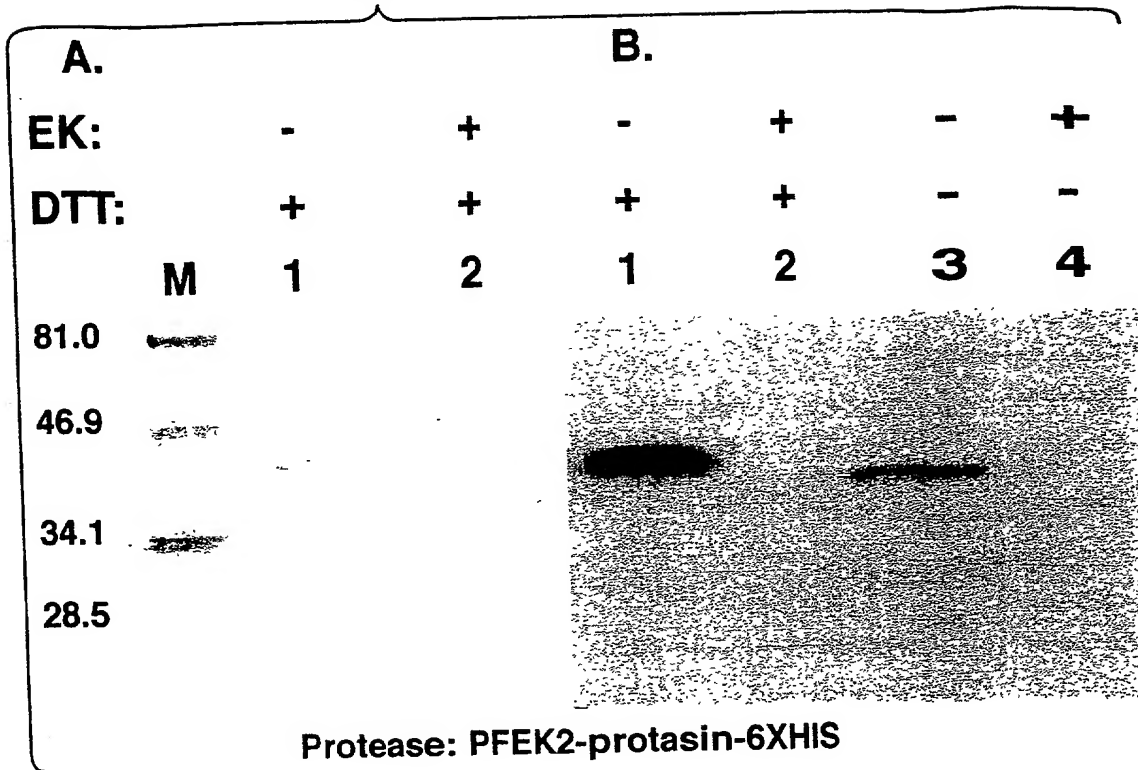
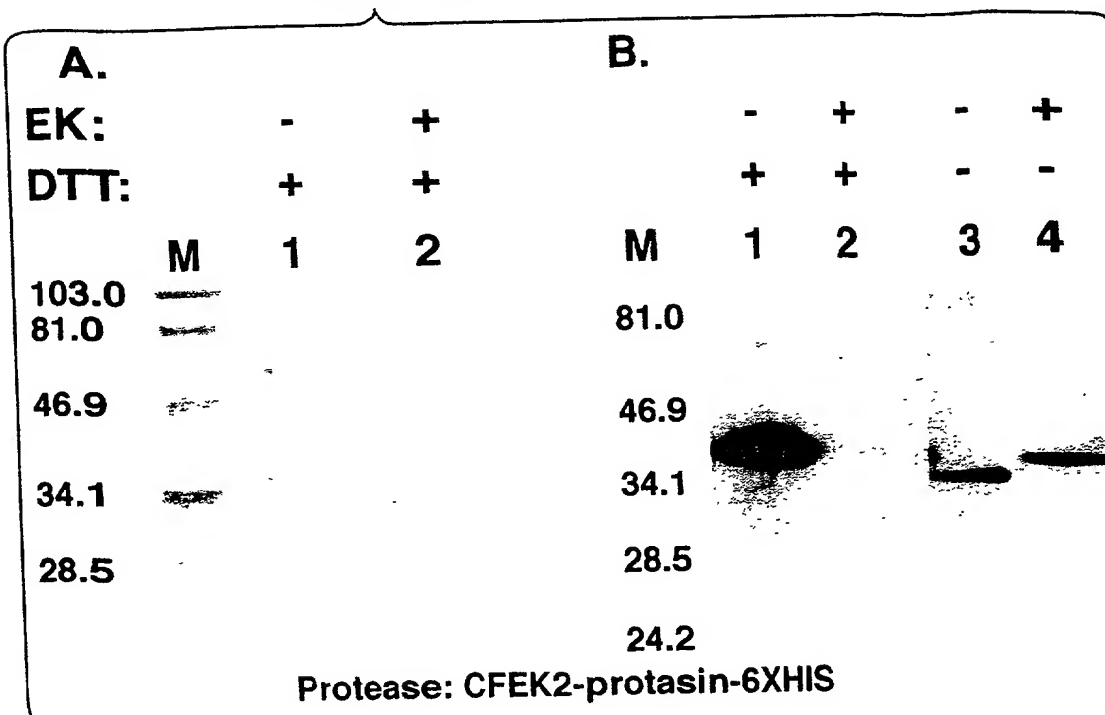
FIG. 7**FIG. 8**

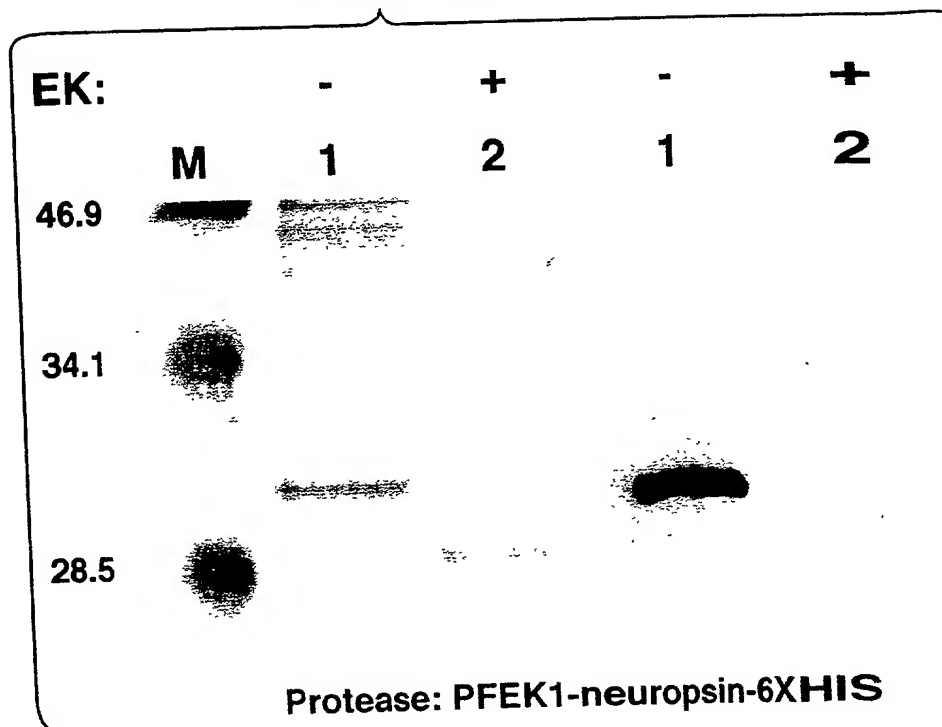
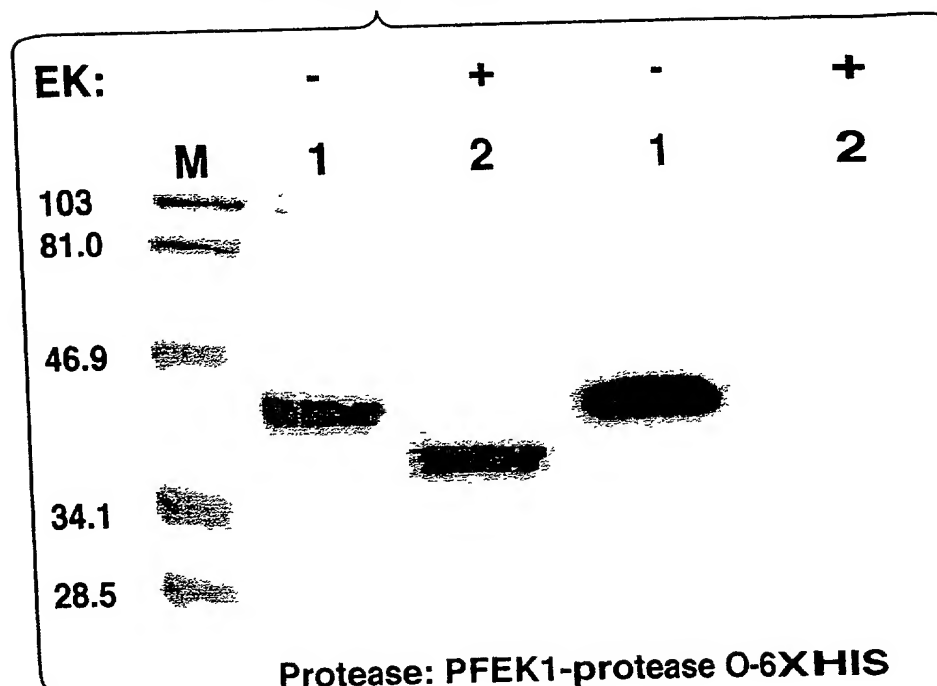
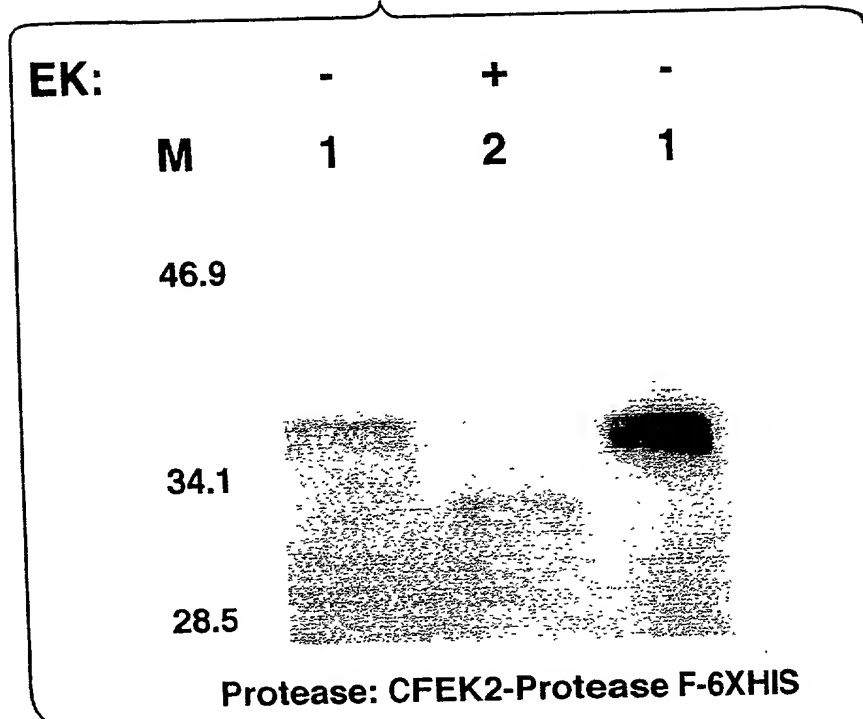
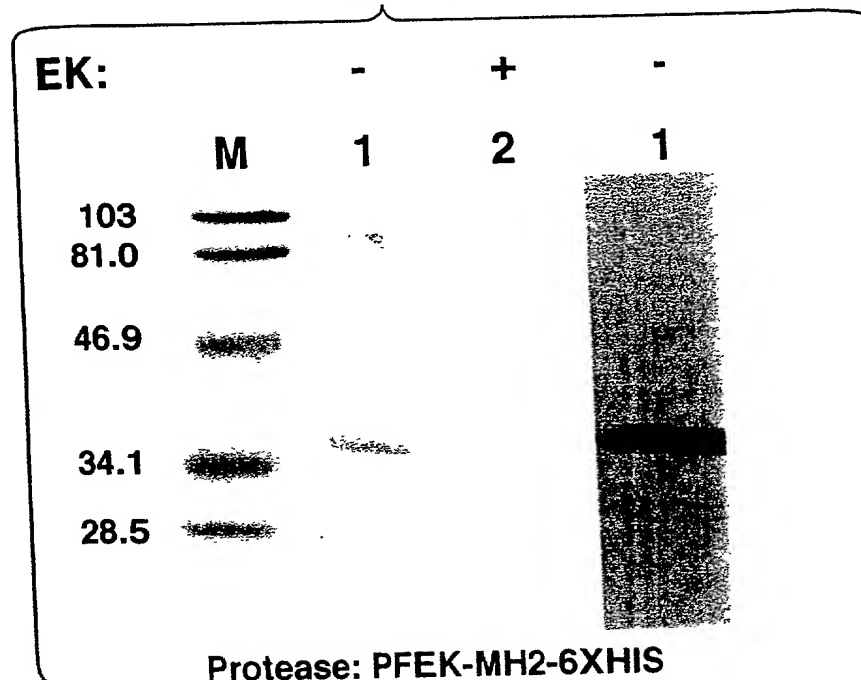
FIG. 9**FIG. 10**

FIG. 11**FIG. 12**

SEQ.ID.NO.:53

FIG. 13(A)

Eco RI
 1 GAATTCACCACCATGGCTTTCCTCTGGCTCCTCTCCTGCTG **GGCCCTCCT** 50
 -----+-----+-----+-----+-----+
 CTTAAGTGGTGGTACCGAAAGGAGACCGAGGAGAGGACGAC **CCGGGAGGA**
 M A F L W L L S C W A I L
 Chymotrypsinogen Pre

51 GGGTACCACCTTCGGCTGCGGGGTCCCCGACTACAAGGACG **ACGACGACG** 100
 -----+-----+-----+-----+-----+
 CCCATGGTGGGAAGCCGACGCCCCAGGGGCTGATGTTCTCTG **TGCTGCTGC**
 G T T F G C G V P D Y K D D D D
 ---Chymotrypsinogen Pre --- FLAG ---

Not I
 101 CGGCCGCTCTTGCTGCCCCCTTTGATGATGATGACAAGATC **GTTGGGGGC** 150
 -----+-----+-----+-----+-----+
 GCCGGCGAGAACGACGGGGGAACTACTACTACTGTTCTAG **CAACCCCGG**
 A A A L A A P F D D D D K I V G G
 --- EK2 Pro ---

Xba I
 151 TATGCTCTAGAACTCGGGCGTTGGCCGTGGCAGGGGAGCCT **GCGCCTGTG** 200
 -----+-----+-----+-----+-----+
 ATACGAGATCTTGAGCCCGCAACCGGCACCGTCCCCTCGGA **CGCGGACAC**
 Y A L E L G R W P W Q G S L R L W
 --- Protease F.CDS ---

201 GGATTCCCACGTATGCGGAGTGAGCCTGCTCAGCCACCGCT **GGGCACTCA** 250
 -----+-----+-----+-----+-----+
 CCTAAGGGTGCATACGCCTCACTCGGACGAGTCGGTGGCGA **CCCGTGTG**
 D S H V C G V S L L S H R W A L
 --- Protease F.CDS ---

251 CGGCGGCGCACTGCTTTGAAACCTATAGTGACCTTAGTGAT **CCCTCCGGG** 300
 -----+-----+-----+-----+-----+
 GCCGCGCGCTGACGAACTTTGGATATCACTGGAATCACT **AGGGAGGCC**
 T A A H C F E T Y S D L S D P S G
 --- Protease F.CDS ---

301 TGGATGGTCCAGTTTGGCCAGCTGACTTCCATGCCATCCTT **CTGGAGCCT** 350
 -----+-----+-----+-----+-----+
 ACCTACCAGGTCAAACCGGTGCACTGAAGGTACGGTAGGA **AGACCTCGGA**
 W M V Q F G Q L T S M P S F W S L
 --- Protease F.CDS ---

FIG. 13(B)

351 GCAGGCCTACTACAACCGTTACTTCGTATCGAATATCTATCTGAGCCCTC 400
 -----+-----+-----+-----+-----+
 CGTCCGGATGATGTTGGCAATGAAGCATAGCTTATAGATAGACTCGGGAG
 Q A Y Y N R Y F V S N I Y L S P
 ----- Protease F.CDS -----

401 GCTACCTGGGGAATTCACCCTATGACATTGCCTTGGTGAAGCTGTCTGCA 450
 -----+-----+-----+-----+-----+
 CGATGGACCCCTTAAGTGGGATACTGTAACGGAACCACTTCGACAGACGT
 R Y L G N S P Y D I A L V K L S A
 ----- Protease F.CDS -----

451 CCTGTCACCTACACTAAACACATCCAGCCCATCTGTCTCCAGGCCTCCAC 500
 -----+-----+-----+-----+-----+
 GGACAGTGGATGTGATTTGTGTAGGTCGGGTAGACAGAGGTCCGGAGGTG
 P V T Y T K H I Q P I C L Q A S T
 ----- Protease F.CDS -----

501 ATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGGGGGTACA 550
 -----+-----+-----+-----+-----+
 TAAACTCAAACCTTTGGCCTGTCTGACGACCCACTGACCGACCCCATGT
 F E F E N R T D C W V T G W G Y
 ----- Protease F.CDS -----

551 TCAAAGAGGATGAGGCACTGCCATCTCCCCACACCCTCCAGGAAGTTCAG 600
 -----+-----+-----+-----+-----+
 AGTTTCTCCTACTCCGTGACGGTAGAGGGGTGTGGGAGGTCTTCAAGTC
 I K E D E A L P S P H T L Q E V Q
 ----- Protease F.CDS -----

601 GTCGCCATCATAAACAACCTCTATGTGCAACCACCTCTTCTCAAGTACAG 650
 -----+-----+-----+-----+-----+
 CAGCGGTAGTATTTGTTGAGATACACGTTGGTGGAGAAGGAGTTTCAATGTC
 V A I I N N S M C N H L F L K Y S
 ----- Protease F.CDS -----

651 TTTCCGCAAGGACATCTTTGGAGACATGGTTTGTGCTGGCAATGCCCAAG 700
 -----+-----+-----+-----+-----+
 AAAGGCGTTCCTGTAGAAACCTCTGTACCAAACACGACCGTTACGGGTTTC
 F R K D I F G D M V C A G N A Q
 ----- Protease F.CDS -----

30/34

FIG. 13(C)

701 GCGGGAAGGATGCCTGCTTCGGTGAAGTGGACCCTTGGCCTGTAAC 750
 -----+-----+-----+-----+-----+
 CGCCCTTCCTACGGACGAAGCCACTGAGTCCACCTGGGAACCGGACATTG
 G G K D A C F G D S G G P L A C N
 ----- Protease F.CDS -----

751 AAGAATGGACTGTGGTATCAGATTGGAGTCGTGAGCTGGGAGTGGGCTG 800
 -----+-----+-----+-----+-----+
 TTCTTACCTGACACCATAGTCTAACCTCAGCACTCGACCCCTCACCCGAC
 K N G L W Y Q I G V V S W G V G C
 ----- Protease F.CDS -----

801 TGGTCGGCCCAATCGGCCCGGTGTCTACACCAATATCAGCCACCACTTG 850
 -----+-----+-----+-----+-----+
 ACCAGCCGGGTAGCCGGGCCACAGATGTGGTTATAGTCGGTGGTGAAC
 G R P N R P G V Y T N I S H H F
 ----- Protease F.CDS -----

851 AGTGGATCCAGAAGCTGATGGCCCAGAGTGGCATGTCCCAGCCAGACCCC 900
 -----+-----+-----+-----+-----+
 TCACCTAGGTCTTCGACTACCGGGTCTCACC GTACAGGGTCGGTCTGGGG
 E W I Q K L M A Q S G M S Q P D P
 ----- Protease F.CDS -----

901 Xba I Not I
 TCCTGGTCTAGACATCACCATCACCATCACTAGCGGCCGCTTCCCTTTAG 950
 -----+-----+-----+-----+-----+
 AGGACCAGATCTGTAGTGGTAGTGGTAGTGATCGCCGGCGAAGGGGAATC
 S W | S R | H H H H H H *
 ----- 6 X HIS-TAG -----

951 TGAGGGTTAATGCTTCGAGCAGACATGATAAGATAATTGATGAGTTGG 1000
 -----+-----+-----+-----+-----+
 ACTCCCAATTACGAAGCTCGTCTGTACTATTCTATGTAAC TACTCAACC

 SV40 Late pA

1001 ACAAAACCACAAGTAGAATGCAGTGAAAAAATGCTTTATTTGTGAAATTT 1050
 -----+-----+-----+-----+-----+
 TGTTTGGTGTGATCTTACGTCACTTTTTTTACGAAATAAACACTTTAA

 SV40 Late pA

1051 GTGATGCTATTGCTTTATTTGTAACCATTATAAGCTGCAATAAACAGTT 1100
 -----+-----+-----+-----+-----+
 CACTACGATAACGAAATAAACATTGGTAATATTCGACGTTATTTGTTCAA

 SV40 Late pA

FIG. 13(D)

1101 GAC 1103

CTG

FIG. 14(A)

SEQ.ID.NO.:54

Eco RI
 1 GAATTCACCACCATGGACAGCAAAGGTTTCGTCGCAGAAATCCCGCTGCT 50
 -----+-----+-----+-----+-----+
 CTTAAGTGGTGGTACCTGTCGTTTCCAAGCAGCGTCTTAGGGCGGACGA
 M D S K G S S Q K S R L L
 Prolactin Signal Sequence

51 CCTGCTGCTGGTGGTGTCAAATCTACTCTTGTGCCAGGGTGTGGTCTCCG 100
 -----+-----+-----+-----+-----+
 GGACGACGACCACCACAGTTTAGATGAGAACACGGTCCCAACACAGAGGC
 L L L V V S N L L L C Q G V V S
 Prolactin Signal Sequence

Not I
 101 ACTACAAGGACGACGACGACGTGGACGCGGCCGCTCTTGCTGCCCCCTTT 150
 -----+-----+-----+-----+-----+
 TGATGTTCTGCTGCTGCTGCACCTGCGCCGGCGAGAACGACGGGGGAAA
 D Y K D D D D V D A A A L A A P F
 FLAG EK1 Pro

Xba I
 151 GATGATGATGACAAGATCGTTGGGGGCTACAACCTGTCTAGAGCCGCACTC 200
 -----+-----+-----+-----+-----+
 CTACTACTACTGTTCTAGCAACCCCGATGTTGACAGATCTCGGCGTGAG
 D D D D K I V G G Y N C L E P H S
 EK1 Pro

201 GCAGCCCTGGCAGGCGGCACTGGTCATGGAAAACGAATTGTTCTGCTCGG 250
 -----+-----+-----+-----+-----+
 CGTCGGGACCGTCCGCCGTGACCAGTACCTTTTGCTTAACAAGACGAGCC
 Q P W Q A A L V M E N E L F C S
 MH2.CDS

251 GCGTCCTGGTGCATCCGCAGTGGGTGCTGTCAGCCGCACACTGTTTCCAG 300
 -----+-----+-----+-----+-----+
 CGCAGGACCACGTAGGCGTCACCCACGACAGTCGGCGTGTGACAAAGGTC
 G V L V H P Q W V L S A A H C F Q
 MH2.CDS

301 AACTCCTACACCATCGGGCTGGGCCTGCACAGTCTTGAGGCCGACCAAGA 350
 -----+-----+-----+-----+-----+
 TTGAGGATGTGGTAGCCCGACCCGACGTGTCAGAACTCCGGCTGGTCT
 N S Y T I G L G L H S L E A D Q E
 MH2.CDS

FIG. 14(B)

351 GCCAGGGAGCCAGATGGTGGAGGCCAGCCTCTCCGTACGGCACCCAGAGT 400
 -----+-----+-----+-----+-----+
 CGGTCCCTCGGTCTACCACCTCCGGTCGGAGAGGCATGCCGTGGGTCTCA
 P G S Q M V E A S L S V R H P E
 ----- MH2.CDS -----

401 ACAACAGACCCTTGCTCGCTAACGACCTCATGCTCATCAAGTTGGACGAA 450
 -----+-----+-----+-----+-----+
 TGTGTCTGGAACGAGCGATTGCTGGAGTACGAGTAGTTCAACCTGCTT
 Y N R P L L A N D L M L I K L D E
 ----- MH2.CDS -----

451 TCCGTGTCCGAGTCTGACACCATCCGGAGCATCAGCATTGCTTCGCACTG 500
 -----+-----+-----+-----+-----+
 AGGCACAGGCTCAGACTGTGGTAGGCCTCGTAGTCGTAACGAAGCGTCAC
 S V S E S D T I R S I S I A S Q C
 ----- MH2.CDS -----

501 CCCTACCGCGGGGAACCTCTTGCCTCGTTTCTGGCTGGGGTCTGCTGGCGA 550
 -----+-----+-----+-----+-----+
 GGGATGGCGCCCTTGAGAACGGAGCAAAGACCGACCCAGACGACCGCT
 P T A G N S C L V S G W G L L A
 ----- MH2.CDS -----

551 ACGGCAGAATGCCTACCGTGCTGCAGTGCGTGAACGTGTCGGTGGTGTCT 600
 -----+-----+-----+-----+-----+
 TGCCGTCTTACGGATGGCAGCAGCTCACGCACTTGCACAGCCACCACAGA
 N G R M P T V L Q C V N V S V V S
 ----- MH2.CDS -----

601 GAGGAGGTCTGCAGTAAGCTCTATGACCCGCTGTACCACCCACAGCATGTT 650
 -----+-----+-----+-----+-----+
 CTCTCCAGACGTCATTTCGAGATACTGGGCGACATGGTGGGGTTCGTACAA
 E E V C S K L Y D P L Y H P S M F
 ----- MH2.CDS -----

651 CTGCGCCGGCGGAGGGCACGACCAGAAGGACTCCTGCAACGGTGACTCTG 700
 -----+-----+-----+-----+-----+
 GACGCGGCCGCTCCCGTGCTGGTCTTCCTGAGGACGTTGCCACTGAGAC
 C A G G G H D Q K D S C N G D S
 ----- MH2.CDS -----

FIG. 14(c)

701 GGGGGCCCCTGATCTGCAACGGGTACTTGCAGGGCCTTGTGTCCTTCGGA 750
 -----+-----+-----+-----+-----+
 CCCCCGGGGACTAGACGTTGCCCATGAACGTCCCGGAACACAGAAAGCCT
 G G P L I C N G Y L Q G L V S F G
 ----- MH2.CDS -----

751 AAAGCCCCGTGTGGCCAAGTTGGCGTGCCAGGTGTCTACACCAACCTTG 800
 -----+-----+-----+-----+-----+
 TTTTCGGGGCACACCGGTTCAACCGCACGGTCCACAGATGTGGTTGGAGAC
 K A P C G Q V G V P G V Y T N L C
 ----- MH2.CDS -----

801 CAAATTCAGTGGATAGAGAAAACCGTCCAGGCCAGTTCTAGACATC 850
 -----+-----+-----+-----+-----+
 GTTTAAGTGAATCACCTATCTCTTTTGGCAGGTCCGGTCAAGATCTGTAG
 K F T E W I E K T V Q A S | S R H
 ----- MH2.CDS -----

851 ACCATCACCATCACTAGCGGCCGCTTCCCTTTAGTGAGGGTTAATGCTTC 900
 -----+-----+-----+-----+-----+
 TGGTAGTGGTAGTGATCGCCGGCGAAGGGAAATCACTCCCAATTACGAAG
 H H H H H *
 ----- 6 X HIS-TAG -----

901 GAGCAGACATGATAAGATACATTGATGAGTTTGGACAAACCACTAGTA 950
 -----+-----+-----+-----+-----+
 CTCGTCTGTACTATTCTATGTAACACTCAAACCTGTTGGTGTGATCT

SV40 Late pA

951 ATGCAGTGAAAAAATGCTTTATTTGTGAAATTTGTGATGCTATTGCTTT 1000
 -----+-----+-----+-----+-----+
 TACGTCACTTTTTTTTACGAAATAAACACTTTAAACACTACGATAACGAAA

SV40 Late pA

1001 ATTTGTAACCATTATAAGCTGCAATAAACAAGTTGAC 1037
 -----+-----+-----+-----+-----+
 TAAACATTGGTAATATTCGACGTTATTTGTTCAACTG